

Data Visualization Guideline

Overview

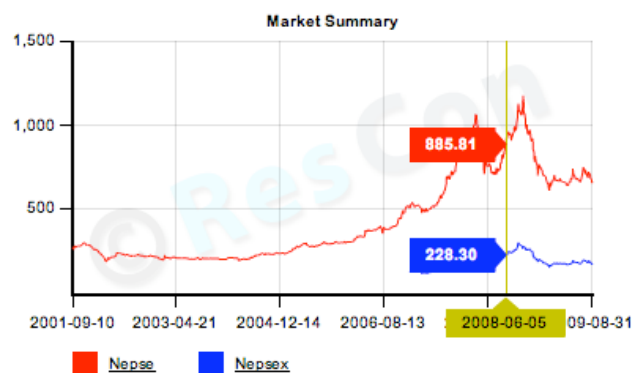
Introduction

Essence of Data visualization is self explanatory, they present the best way to answer questions. Whenever the set of data is produced in form of table or a list it contains information, data visualization is the method to extract, preset aesthetically and logically the data in human readable form, same data-set can produce different visualization according to particular question asked.

visualization types and usages

| Charts | Graphs | Special Visualization |
|--|--|---|
| Common bar-charts, pie-charts, scatter diagram, line/area chart etc. | Venn diagram, flowchart, UML diagrams, connected networks, cluster | Fusion of chart and graphs, special mathematical visualizations |
| Simple data-set, value changes, dispersion, skewness etc. | Interference, processing, influences, footprint, effects etc. | Representation of large data-set, special inferencing, aesthetic plots etc. |

Appropriate usages of these visualization techniques largely depend upon questions the diagram is answering and target audience, Well represented diagrams serves multiple purpose. It adds aesthetic to whole article, increase readership, provide a glance inside news, serves as future reference and boost quality of newspaper at once.



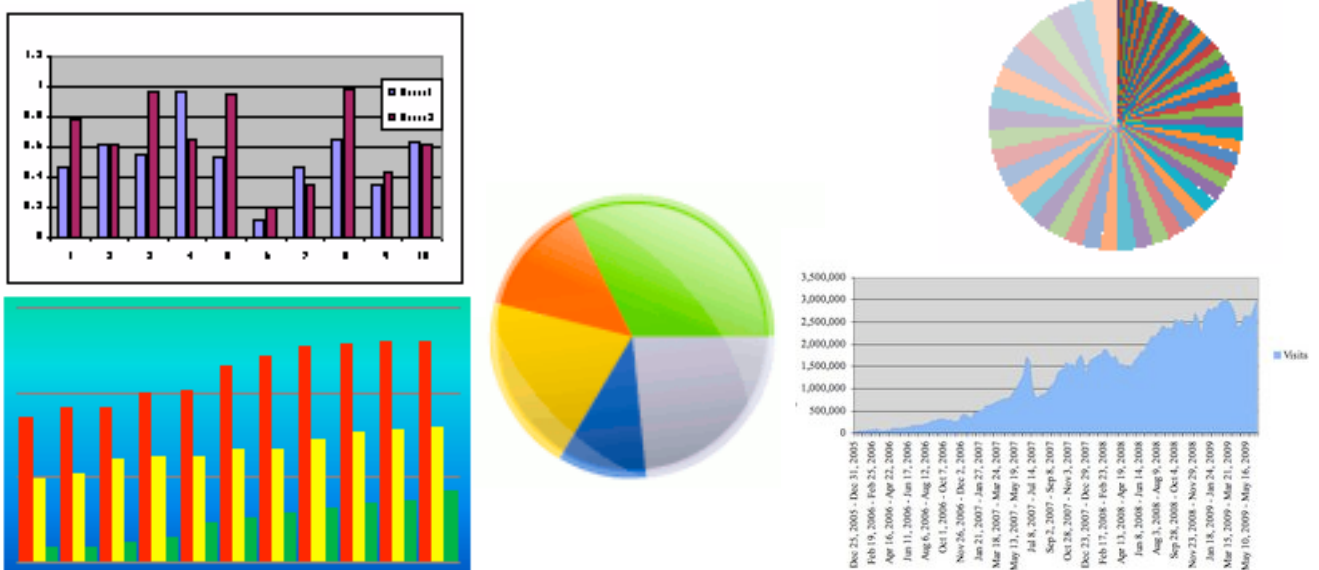
Market summary of Nepse from <http://www.res-con.biz/>

In this Document

- This document is brief summary of:
 - Visualization guideline
 - Tools for visualizing data
 - References

Guideline

- **Start with Data**, don't start with chart type. Spot the distinct trend in data and visualize what is important and how it can be conveyed better, after finding trend in data it is easier and effective to choose chart type.
- **Use available data wisely**, it is not necessary to plot all data at once. There are options for multiple charts and you can choose over certain frequency periods to highlight the message. Better option is to ask what question the visualization should answer and choose number and type of visualization.
- **Don't start with bar/pie chart**, Pie chart have it's own purpose (showing part to whole ratio) and charts are not always rectangular bar to begin with. If only bar and pie charts comes as option then there is some flaw in thinking or data itself. Perhaps learning about different types of charts helps a lot
- **Never use default template**, there are many reason not to use default layouts. Primarily they are overused and thus monotonous, secondarily they are templates which might not be suitable to particular data-set. Similarly using templates without modifying makes diagram maker look lazy and out of taste.
- **Avoid primary colors**, Primary color are hard over the eyes. Primary color catches attention easily and if multiple primary colors are used the charts are hard to read, use gray shades for grid (and background if necessary) and dull colors for data, if required use 1-2 primary colors to highlight particular point on chart
- **Never skew/bias the charts**, even if it is beneficial. It is easy to bias the charts, ever minor changes in scale and frequency of the plot produces different inferences, but once you are caught biasing whole reputation is gone.
- **Charts are not interchangeable**, sometime data-set may be compatible with more than one charts, but data conveys different message in different charts, so never start with one chart and later change it to other, do careful study beforehand .
- **Don't design charts individually**, charts alone are less meaningful. Charts are most of the time used with document, news, report etc. which have distinctive layout, color combination and presentation style on overall. Charts must fit within. If overall layout is not considered while designing charts, they will not aid the purpose at all.
- **Everything 3d is not good**, while aesthetic is necessary to attract the audience, it does not mean that eye-candy and 3d charts should be overused, if simple charts conveys message well and are pleasing to eyes they should be opted over 3d and candy charts.



Few charts displaying fundamental flaws in representation

Recommended Tools

| Visualization | Tools | |
|--|--|---|
| Bar charts, pie charts, line charts, scatter charts, area charts, Venn diagram, organization charts, simple network graphs. (for daily usages and rapid development) | Apps | Remarks |
| | Microsoft Excel | Basic graphs, low customization |
| | Apple Numbers | Works with OS X only, basic graphs |
| Visualization tools with Graphical user interface and tools requiring medium level statistical and mathematical skills. (for regular usages with few hours to play over data-set) | Adobe Illustrator | Basic graphs, highly customizable |
| | Apps | Remarks |
| | OpenDX | OSX version costs \$25, windows version is free. |
| | Weka | data mining tool, free of cost, requires some learning. |
| | Microsoft Visio | drawing and diagramming software from microsoft. |
| Visualization tools with Programmable IDE and requiring programming and mathematical skills. (for special usages with atleast 1-2 days to play over data-set) | Crystal Xcelcius | Visualization solution from Crystal, can also be included in webpages |
| | Apps | Remarks |
| | Axiis | Open source technology build over adobe flex |
| | Google Visualization API | Google's answer for visualization (require programming skills) |
| | IBM ManyEyes | IBM's intuitive visualization API |

References

<http://www.gapminder.org/>

Gapminder is a non-profit venture promoting sustainable global development and achievement of the United Nations Millennium Development Goals by increased use and understanding of statistics and other information about social, economic and environmental development at local, national and global levels.

<http://www.gnuplot.info/>

Gnuplot supports many types of plots in either 2D and 3D. It can draw using lines, points, boxes, contours, vector fields, surfaces, and various associated text. It also supports various specialized plot types

<http://opendx.org/>

With OpenDX, you can create the visualizations you want to create. OpenDX has been designed to be the place where the art of science and the science of visualization come together. It's the place where they're combined into one powerful, flexible framework that lets you "Simply Visualize."

<http://www.cs.waikato.ac.nz/ml/weka/>

Weka is a collection of machine learning algorithms for data mining tasks. The algorithms can either be applied directly to a dataset or called from your own Java code. Weka contains tools for data pre-processing, classification, regression, clustering, association rules, and visualization. It is also well-suited for developing new machine learning schemes.

<http://www.stata.com/>

Stata is a general-purpose statistical software package created in 1985 by StataCorp. It is used by many businesses and academic institutions around the world. Most of its users work in research, especially in the fields of economics, sociology, political science, and epidemiology.

<http://www.datavisualization.ch/>

Datavisualization.ch is a website dedicated to data and information visualization. The mission is to document and discuss research findings in this field. This includes cognitions from self initiated studies as well as providing an overview of the development done by the incredible smart people in the community.

<http://research.microsoft.com/en-us/groups/dmx/>

Microsoft Data Visualization center,

<http://www.in-rev.com>

InRev Systems Bangalore Pvt. Ltd transforms business by making all possible information available at all times. InRev combines domain knowledge, trusted reporting structure and innovative applications to help companies advance their business decisions. InRev serves across all stages of Business Intelligence and Decision Management right from Data Extraction and Integration to Strategy Consulting using the intelligent insights of the business. InRev is the promise to the World to take the Information Extraction and Usage to the next level. InRev plans to make use of every possible data and process them to build Killer Apps which would bring an Information Revolution in the World.

<http://www.axiis.org/about.html>

Axiis is an open source data visualization framework designed for beginner and expert developers alike.

Whether you are building elegant charts for executive briefings or exploring the boundaries of advanced data visualization research, Axiis has something for you.

Axiis provides both pre-built visualization components as well as abstract layout patterns and rendering classes that allow you to create your own unique visualizations.

<http://www.satimage-software.com/>

Smile is a Macintosh computer programming and working environment based on AppleScript. It features a number of production technologies and a natural fashion of having them work together. Smile is primarily designed for scientists, engineers, desktop publishers, and web administrators, to help them producing faster and better, automating frequent tasks, and controlling complex operations.

<http://www.res-con.biz/>

ResCon, pronounced as /res'kun/, is the brainchild of Anup Sharma and Tilak Acharya. After closely monitoring the financial sector, Anup and Tilak felt the need of a company providing cognitive analysis of the Nepalese capital market. They observed that many customers were willing to invest but didn't have a beacon light showing the best possible direction to invest in the Nepalese Stock Market. As both of them had good background knowledge of the Nepalese Stock Market, they were confident of bridging this gap and providing the "beacon light" for the prospective investors. Hence, with a clear set of vision, they started gathering data and analyzing them since the Spring of 2007. After almost one year of constant research, the idea finally took the shape and culminated into ResCon, an amalgamation of Research and Consulting.

<http://vizlab.nytimes.com/>

Visualization Lab, where you can create visual representations of data and information using the "Many Eyes" technology from IBM Research.

<http://code.google.com/apis/visualization/>

The Google Visualization API lets you access multiple sources of structured data that you can display, choosing from a large selection of visualizations. Google Visualization API enables you to expose your own data, stored on any data-store that is connected to the web, as a Visualization compliant datasource. Thus you can create reports and dashboards as well as analyze and display your data through the wealth of available visualization applications. The Google Visualization API also provides a platform that can be used to create, share and reuse visualizations written by the developer community at large.

<http://manyeyes.alphaworks.ibm.com/manyeyes/>

Many Eyes is an IBM Research project and website whose stated goal is to democratize information and to enable social data analysis ("social" in the sense of Web 2.0), by making it easy for laypeople to create, edit, share and discuss each other's information visualizations. It was started in 2007 by Fernanda Viégas and Martin Wattenberg.

This work is licensed under the Creative Commons Attribution-Share Alike 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-sa/3.0/> or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

